Huaying Gu

ON, Toronto| 0086-13942020915| hy491291367@gmail.com|
Linkedin www.linkedin.com/in/huaying-gu-3636a3239
Portfolio https://kkkhora.github.io/IT-job-portfolio/

_____

SUMMARY

An IT fresher aspires to become a Software Developer looking for an internship in Software Engineering, with knowledge covering both Computer and Data science.
- A multidisciplinary professional with a major in both Art and Science.
- A quick learner, excited to learn all kinds of new skills and knowledge in all subjects.
- Good at team collaboration, and time management, and love to deal with challenges.

TECHNICAL SKILLS

Operating System: Windows, Mac, Linux
Programming Language: java, python, JavaScript, SQL, C, C++
Database: Oracle, MySQL, PostgreSQL, Mango DB, Neo4j
Other Technology: HTML, CSS, Node.JS, Express.JS, EJS, REST API, React, React Hooks BootStrap, GIT, Apache Spark, Tensorflow

PROJECTS

SPOTIHOLIC – Spotify Recommendation Web App   **Feb 2022 – Apr 2022**
Group project (4 people) - Database and Information System, achieved a high grade of 97/100
- Search and learn how to use public API and lead the team to collect data of over 24000 songs and their information using Spotify and Musixmatch API, laying a solid foundation for the whole project to start.
- Supervise the team in designing and building API such as searching songs in different ways, user login/register, playlist display, recommendation, etc. solving the most crucial part of the project, and successfully completing the main functions based on it.
- Generate an interactive module with a world map for users to select songs based on country. This user-interact function exceeded the basic requirement of the assignment and made the application more intelligent.
- Optimize SQL queries with indexing, materialized view and halve the execution time.

Big Data Analysis on Wine Review   **Feb 2023 – Apr 2023**
Kaggle competition-styled group project (3 people) - Big Data Analytics, 100% result of 105/105
- Wrangle, transform data from a wine review file with over 280000 pieces, and extract Key information, making the overall data cleaner and more readable.
- Use Natural Language Toolkit (nltk) to process natural language, including tokenizing the "description" strings, creating a pos tag from the tokens, doing sentiment analysis upon the reviews, and generating a word cloud. This serves as a crucial process in predicting the wine price, as the review of wine is one of the most significant features.
- Construct a neural network model using PyTorch to predict the price of a wine, completing the most complex supervised learning model in this assignment.

Division by Zero Debugger   **Jun 2022 – Aug 2022**
Group project (2 people) - Software Analysis, high grade of 99/100
- Design code structure using static analysis in particular dataflow analysis and pointer analysis, and draw the project architecture, making the project clear to implement.
- Define abstract domain for Nonvalue, NonNull, Maybe Null, Null, making the debugger identify the state of a variable more precisely beyond simply Null and NonNull.
- Implement Chaotic Iteration algorithm using the following helper function: getPredecessors, join, flowIn, flow-out, and "transfer" function focusing on the following instructions: Alloca,

    Store, Load, Call, Cmp, BinOp, fixing several Key bugs to make the debugger function well with no error.
- Design different test cases to calculate Precision, Recall, and F-measure, making the test cases more complicated: not only include simple non-null/null pointers but also branches, for-loops, double pointers, uninitialized pointers, malloc memory allocation, etc. This optimization made the debugger more powerful in identifying different and complex div by zero cases.

Covid Data Processor   **Oct 2021 – Nov 2021**
Group project (2 people) Data Structure and Software Design, received full grades of 100/100
- Process the given data of over 10000 local COVID test, vaccination, hospitalization, and death cases, over 580000 local housing properties, and the population of each location, successfully sort large data by choosing the right data structure to implement.
- Optimize the performance using efficient data structures such as Tree Map for ordering the Zip Code and LinkedList for cross-referencing, improving the performance by 20%.

EDUCATION
Master in Computer and Information Technology   2021- 2025- In Process
University of Pennsylvania, USA

Master in Gender Studies   2018 graduated
University of Cambridge, UK

Bachelor in Art History   2016 graduated
Peking University, China